

# Nonparametric forecasting outperforms parametric methods for a simulated multispecies system

CHARLES T. PERRETTI,<sup>1,3</sup> GEORGE SUGIHARA,<sup>1</sup> AND STEPHAN B. MUNCH<sup>2</sup>

<sup>1</sup>*Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California 92093 USA*

<sup>2</sup>*Fisheries Ecology Division, Southwest Fisheries Science Center, 110 Shaffer Road, Santa Cruz, California 95066 USA*

**Abstract.** Ecosystem dynamics are often complex, nonlinear, and characterized by critical thresholds or phase shifts. To implement sustainable management plans, resource managers need to accurately forecast species abundance. Moreover, an ecosystem-based approach to management requires forecasting the dynamics of all relevant species and the ability to anticipate indirect effects of management decisions. It is therefore crucial to determine which forecasting methods are most robust to observational and structural uncertainty. Here we describe a nonparametric method for multispecies forecasting and evaluate its performance relative to a suite of parametric models. We found that, in the presence of noise, it is often possible to obtain more accurate forecasts from the nonparametric method than from the model that was used to generate the data. The inclusion of data from additional species yielded a large improvement for the nonparametric model, a smaller improvement for the control model, and only a slight improvement for the alternative parametric models. These results suggest that flexible nonparametric modeling should be considered for ecosystem management.

**Key words:** ecosystem-based management; ecosystem modeling; forecasting; nonlinear dynamics.

## INTRODUCTION

To effectively manage a species, one must have a method for accurately describing its dynamics. Moreover, the call for ecosystem-based management requires a method for simultaneously predicting the dynamics of many species (Grumbine 1994, Slocumbe 1998, Pikitch et al. 2004). This task is complicated by ecosystem dynamics that often are complex, nonlinear, and exhibit phase shifts; while our understanding of them is usually derived from short, noisy time series of just a few variables. Therefore, it is important to evaluate which forecasting methods are most robust to uncertainty and to extend these methods for use in conservation and management.

Many ecological modeling frameworks are available for forecasting, ranging from simple single-species models to highly complex ecosystem models (Schaefer 1957, Christensen and Walters 2004, Fulton et al. 2005). Recently, there has been a trend toward developing increasingly complex models in hopes of increasing model realism. However, increased model complexity will not necessarily lead to increased forecast accuracy, as network topology and the functional relationships between species are often highly uncertain and model identifiability is a major challenge (Ludwig et al. 1988). Moreover, apparently small changes in model structure

may produce very large, qualitative changes in predictions (Wood and Thomas 1999). Nonparametric forecasting methods (e.g., Hardle et al. 1997), and particularly state-space reconstruction techniques (SSR; Sugihara 1994), are more robust to such structural uncertainty and may provide a way forward.

Past studies have used SSR to identify chaos and, more generally, nonlinearities in ecological time series (Sugihara et al. 1990, Hsieh et al. 2005), and although SSR techniques have gained wide popularity in other fields (Schreiber 1999, Sugihara and May 1990), they have not been used extensively by natural resource managers (but see Glaser et al. 2011). One reason is that, as traditionally applied, SSR methods require long, highly precise time series, which are rare in ecology. To overcome this constraint, multivariate SSR techniques have been developed, as well as methods for compositing time series (Dixon et al. 1999, Hsieh et al. 2008, Deyle and Sugihara 2011; H. Ye et al., *unpublished manuscript*).

Here, we explicitly describe a method for nonparametric multivariate SSR and forecasting that we call the MS-Map (MultiSpecies Map), and evaluate its performance relative to a suite of parametric models fit to simulated data.

## MULTIVARIATE TIME DELAY EMBEDDING

Taken's theorem of state-space reconstruction (Taken 1981) allows for the representation of a multivariate dynamical system through time delay embedding. As traditionally applied, a univariate time series is transformed into a set of time-delayed vectors:

Manuscript received 31 May 2012; revised 7 November 2012, accepted 10 December 2012. Corresponding Editor: S. J. Schreiber.

<sup>3</sup> E-mail: cperrett@ucsd.edu

$$\mathbf{X}_t = [x_t, x_{t-\tau}, x_{t-2\tau}, \dots, x_{t-(E-1)\tau}]$$

where  $x$  is a scalar time series,  $t$  is the time index,  $\tau$  is the time lag, and  $E$  is the embedding dimension. Provided that  $E$  is sufficiently large, the collection of vectors  $\mathbf{X}_t$ , for  $t = [1 + (E - 1)\tau, \dots, N]$ , is an embedding of the attracting manifold, where  $N$  is the length of the time series. Using this procedure, we generate an attractor for each species, then combine that information through a multivariate smoothing kernel we refer to as the MS-Map.

Predictions using the MS-Map are made by first dividing the time-delayed points into a training set and a test set. The length of the training set was varied from 45 to 75 points, and the length of the test set was fixed at 60 points (Appendix: Table A1). The  $p$ -step-ahead prediction for each test point is the distance-weighted average of the positions of all training points  $p$ -steps forward in time. Thus, similar to previous work (Sugihara 1994, Sugihara et al. 1999), predictions of the MS-Map are locally weighted means with weights determined by the distances in delay coordinates. The weight of each training point is recalculated for each test point as follows:

$$w_{ij} = \exp \left[ -\sum_k \frac{\theta_k d_{ijk}}{\bar{d}_j} \right].$$

In the univariate case,  $k = 1$ , and the weights are simply determined by the Euclidean distance,  $d_{ij}$ , between delay coordinates of the  $i$ th training point and the  $j$ th test point. The parameter  $\theta_k$  describes the degree of local weighting, and  $\bar{d}_j$  is the average distance between test point  $j$  and the training points.

An obvious extension of the univariate case to multispecies data would be to use Euclidean distances among the multivariate delay coordinates. However, that approach implicitly treats all species equally in determining predictions, which may be inappropriate in cases where abundance scales are dramatically different or only a subset of the species are important. Therefore, the MS-Map allows each species in the system to affect the weight parameter independently and not necessarily equally by adaptively choosing  $\theta_k$  separately for all species. For example, if  $\theta_k = 0$ , then species  $k$  does not contribute to the forecast, while  $\theta_k > 0$  gives species  $k$  more weight. The  $p$ -step-ahead prediction for each test point is given by

$$\hat{y}_j = \frac{1}{\sum_i w_{ij}} \sum_i w_{ij} x_{i+p}$$

where  $w_{ij}$  is the weight of training point  $i$  for test point  $j$ , and  $x_{i+p}$  is the  $p$ -step-ahead value of training point  $i$  on the reconstructed attractor. A similar method was developed by Cao et al. 1998, and applied to several physical systems; however, unlike our method, equal weights were given to all variables.

To limit the computational requirements of the parameter search for this expository analysis, we restricted the attractor reconstructions to a time lag of unity and an embedding dimension of three. This represents a conservative evaluation of the method as the best embedding dimension may be up to  $2n + 1$  (Takens 1981). To avoid over-fitting, the last 15 points of the training set were excluded and used to determine the values of  $\theta$  that resulted in the lowest forecast error (alternatively this could be done using cross-validation, e.g., Schreiber 1999). Those parameters were then used to forecast the out-of-sample test points. The time series of each species was scaled to  $[0, 1]$  in order to standardize distances between points in phase space across species. All MS-Map simulations were performed using custom code written in MATLAB version 7.7.0 (see the Supplement for code; MathWorks 2008).

#### PARAMETRIC MODELS AND SIMULATION METHODS

Our interest is in comparing forecast performance for dynamics that are sufficiently complex and data that are sufficiently noisy to be ecologically relevant, while simultaneously amenable to fairly intense simulation. Therefore, we chose the three-species model developed by Hastings and Powell 1991 with added harvesting, given by the following:

$$\frac{dx}{dt} = x(1 - x) - f_1(x)y - Fx$$

$$\frac{dy}{dt} = f_1(x)y - f_2(y)z - d_1y$$

$$\frac{dz}{dt} = f_2(y)z - d_2z$$

$$f_i(u) = \frac{a_i u}{1 + b_i u}$$

where  $x$ ,  $y$ , and  $z$  are abundances of each species  $x$ ,  $y$ , and  $z$ , respectively. Parameters  $a_i$  and  $b_i$  govern the feeding rate (functional response) on species  $i$ , and  $d_i$  is the mortality rate of species  $i$ . The bottom species in the food chain (species  $x$ ) is harvested at a constant rate (parameter  $F$ ), which is analogous to a forage-fish fishery. The model was numerically solved using a fourth-order Runge-Kutta method with a fixed time step. One thousand parameter sets were drawn uniformly from the ranges given in the Appendix: Table A1. We focused on parameter sets that produced dynamics ranging from limit cycles to chaos, excluding steady-state solutions (such as species extinctions), resulting in 236 limit cycle parameter sets and 81 chaotic parameter sets (largest Lyapunov exponent  $> 0.01$ ). We ran the model for an initial period of 50 000 time steps to exclude transient dynamics. To model observation error, we multiplied the time series by log-normal noise for the range of standard deviations listed in the Appendix:

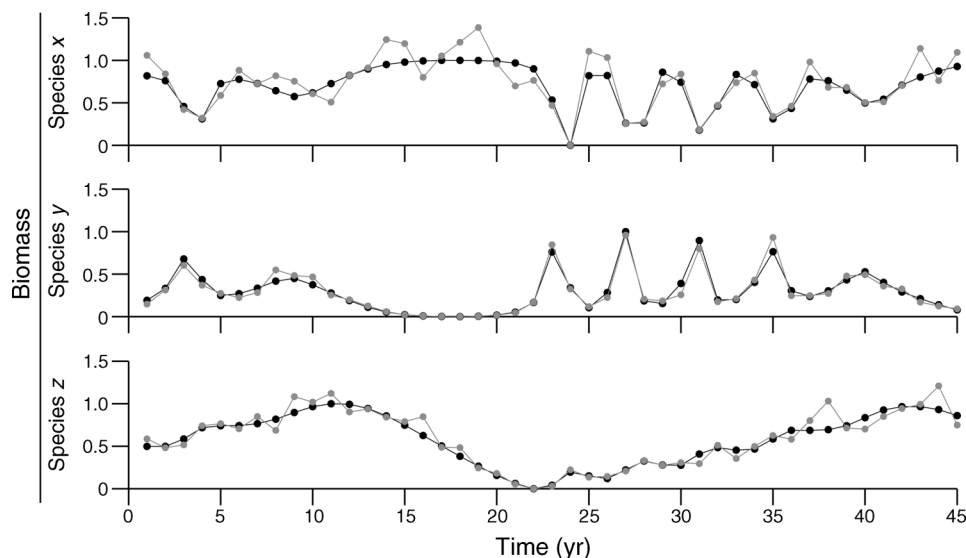


FIG. 1. A typical times series for the short (45-year) simulations of the three-species Hastings-Powell (control) model. The high-noise time series ( $CV = 0.7$ ) are shown in gray, and the deterministic time series are shown in black.

Table A1. The highest noise intensity resulted in a coefficient of variation of each species time series of  $\sim 0.7$ , which is similar to that found in field studies (Fig. 1; Reed and Hobbs 2004).

We focused on forecasting the abundance of the harvested species, which in these simulations was always species  $x$ . We compared the forecast accuracy of the MS-Map algorithm to a single-species Schaefer model:

$$\frac{dx}{dt} = rx \left( 1 - \frac{x}{K} \right) - Fx$$

(Schaefer 1957), a three-species Lotka-Volterra model (LV3; Takeuchi 1996), and the Hastings and Powell model (HP) that was used to generate the data.

Recently, there has been a dramatic increase in the use of models that incorporate both process error and observation error (Cressie et al. 2009). By including process error, it has been suggested that these models are less affected by model misspecification (Clark and Bjørnstad 2004). Simultaneously, there has been a rise in the use of complex ecosystem models (Fulton et al. 2011). Therefore, we also included a more complicated model, a five-species, discrete Lotka-Volterra model that incorporates log-normal process and observation error (LV5):

$$x_i(t+1)$$

$$= x_i(t) \exp \left\{ \varepsilon_i(t) + r_i \left[ 1 - \frac{x_i(t) + \frac{\alpha}{n-1} \sum_{j \neq i}^n x_j(t)}{K_i} \right] \right\}$$

where  $x_i(t)$  is the biomass of species  $i$  at time  $t$ ,  $\varepsilon$  is a normally distributed random variable with mean zero,  $n$

is the number of species, and  $\alpha$  is the competition coefficient which is scaled by  $1/(n-1)$  (Ives et al. 1999).

In the three-species Lotka-Volterra model, the Type II functional response of the HP model is replaced with a Type I functional response, i.e.,  $f_i(u) = a_i u$ . Ecosystem models are frequently linear and bilinear (Christensen and Walters 2004, Steele and Ruzicka 2011), as are statistical models used to describe ecosystems (Holmes et al. 2007, Ives et al. 2010); therefore, this model was chosen to evaluate the ability of a bilinear model to forecast more complex dynamics and to assess the change in forecast accuracy due to an apparently small difference in model structure.

The observed time series was generated by sampling the continuous time series every 50 time steps. The initial conditions and all parameters other than  $F$  (which is assumed to be known) were fit to the observed time series using maximum likelihood for the observation-error models. The five-species state-space model (LV5) was fit using adaptive Bayesian MCMC with multiple chains (see the Supplement for code; Geyer 1991) using R software version 2.14.1 (R Development Core Team 2011), and JAGS version 3.2.0 (Plummer 2003). The variance of the process error and the observation error were both estimated. The parameters used to start the MCMC are given in Appendix: Table A3. We terminated the MCMC routine when the Gelman-Rubin convergence diagnostic was within one percent of unity for all parameters (see Appendix: Fig. A1 for a typical posterior probability plot).

Fitting nonlinear models to noisy data is notoriously difficult (Polansky et al. 2009); therefore, to increase the probability that the global maximum of the likelihood function would be found we initiated the fitting procedure at the correct initial conditions, and for the

HP and LV models, the correct model parameters. Without starting the parameter search near the true values the probability of finding the global maximum degrades (see e.g., Berliner 1991). Importantly, a real practitioner has no guarantee that she tested the “true” parameters. Therefore, our results should be viewed as a best case scenario for the parametric models.

It is well known that long-term predictions for chaotic systems will suffer due to sensitive dependence on initial conditions and non-stationarity. However, it may be possible to make skillful short-term predictions that can then be used for adaptive management. Therefore, we chose to evaluate the change in forecast accuracy of three-step-ahead predictions (which is beyond the autocorrelation threshold) when fitting the models with data from one, two, and three species, and for a range of training set lengths, noise intensities, and harvest rates. The biological meaning of three-step ahead predictions will vary by species and system.

Due to computational constraints, the LV5 model (which includes process and observation error) was only fit using data from all three species. Predictions for the LV5 model were made using the deterministic skeleton, with parameter values corresponding to the mode of the posterior distribution.

To compare performance across models we log-transformed the forecast root mean-square errors (RMSE) for each set of simulations and performed paired *t* tests for all model pairs. The forecasted species was always species *x*, as similar results were found when forecasting species *y* or *z*. We refer to training set lengths of 45, 60, and 75 points as short, medium, and long time series, respectively. Low, medium, and high noise intensities refer to a standard deviation of the normal distribution (used to generate log-normal observation errors) of 0.05, 0.1, and 0.15. Harvest rates of 0.025, 0.05, and 0.075 correspond to low, medium, and high harvest rates, respectively (Appendix: Table A1).

## RESULTS

In the simulations without noise, the HP model generated perfect forecasts, even when fit with only one species time series. This was unsurprising given that the parameter search for the HP model was initiated at the correct values. The second best performing model in the noiseless simulations was the MS-Map with an average error of 13%, 11%, and 10% when fit using one, two, or three species, respectively. In contrast to previous work suggesting simple mechanistic models can outperform the correct mechanistic model (Ludwig and Walters 1985), we found that the Schaefer model performed much worse than the HP model. Furthermore, the Schaefer, LV3, and LV5 models fit to three species all performed substantially worse than the MS-Map, with an average error of 41%, 38%, and 41%, respectively, which was not significantly better than using the mean of the time series as the prediction.

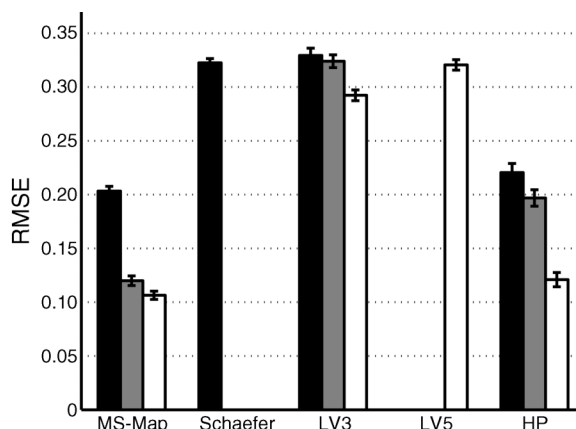


FIG. 2. Average three-step-ahead root mean-square errors (RMSE) of all forecasting methods for all simulations with noise. MS-Map is the multispecies map, LV3 is the Lotka-Volterra three-species model, LV5 is the Lotka-Volterra five-species model, and HP is the Hastings-Powell (control) model. Black, gray, and white bars are the model fit using time series from one, two, and three species, respectively. Error bars show  $\pm$ SE.

Surprisingly, when averaged over all simulations with noise, the MS-Map was more accurate than the HP model by 2.8 percentage points (pp) when fit using one species, 11.3 pp when using two species, and by 1.4 pp when all three species were used; all differences were significant at the  $P < 0.0001$  level (Fig. 2). The difference was most dramatic under the high-noise simulations, where the MS-Map outperformed the HP model by 5.6 pp, 15.5 pp, and 4.3 pp, when fit with one, two, and three species, respectively. Under the low-noise simulations, the MS-Map only outperformed the HP model when fit with two species (5.6 pp,  $P < 0.001$  level). These results were consistent across chaotic and limit cycle parameter sets (Appendix: Fig. A2).

Overall, despite the apparent structural similarity between the LV models and the HP model, the LV models failed to outperform the univariate MS-Map even when the LV models were fit to all three-species time series. The forecast error for both LV models was always at least 38% (Appendix: Table A2).

The MS-Map forecast accuracy improved significantly when additional species were used to fit the model. Across all simulations, there was a 11.3 pp improvement when the second species was added and an additional 2.8 pp improvement when the third species was added ( $P < 0.0001$ ). The HP model also improved with each additional species, with a 2.8 pp and 12.7 pp improvement when the second and third species were added, respectively ( $P < 0.0001$ ). The LV3 model did not show a significant improvement when going from a one- to two-species fit; however, there was a 4.3 pp improvement when moving from one to three species ( $P < 0.001$ ). Increasing the harvest rate resulted in a small improvement in forecast accuracy for the MS-Map



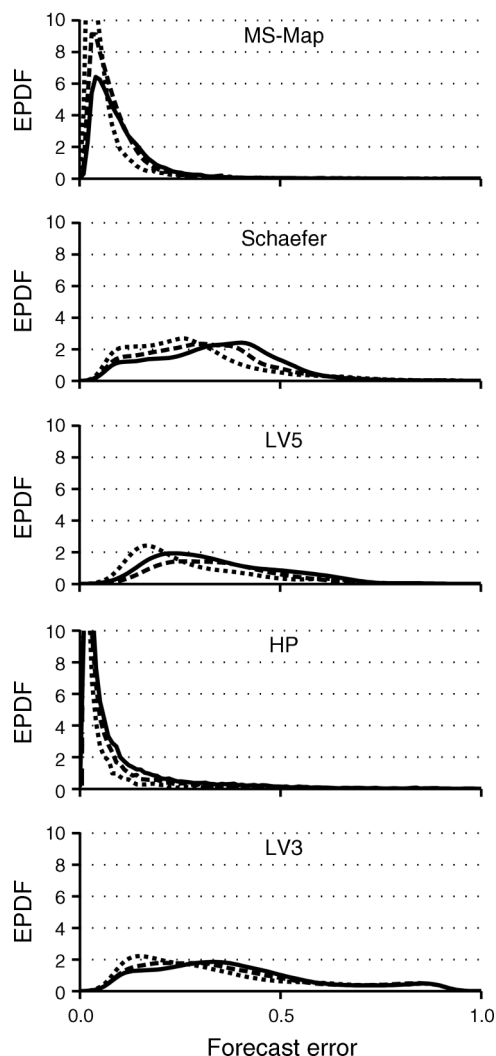


FIG. 3. Empirical probability density function (EPDF) of forecast errors ( $|\text{Predicted} - \text{Actual}|$ ) for each forecasting method. The dotted lines show the low-noise simulations, dashed lines show medium-noise simulations, and solid lines show high-noise simulations. MS-Map is the multispecies map, LV3 is the Lotka-Volterra model, LV5 is the Lotka-Volterra five-species model, and HP is the Hastings-Powell (control) model.

model; however, the effect on the parametric models was mixed and not statistically significant.

#### DISCUSSION

The forecast performance of the nonparametric method was generally superior to that of the parametric models. For both chaotic and limit cycle parameter sets, the MS-Map always produced more accurate forecasts than the incorrectly specified parametric models, and frequently outperformed the correct parametric model (Fig. 2 and Appendix: Fig. A2). These results show that, in the presence of noise, it is often possible to obtain more accurate forecasts from a nonparametric model

than from the model that was used to generate the data. One possible reason for this counterintuitive result is that the likelihood surface of the HP model is extremely rugose, as small changes in the parameters or initial conditions produce large changes in dynamics, especially when the dynamics are near chaotic. This rugosity combined with high observation error causes the best-fit parameters to differ from the true parameters.

Importantly, we found that a relatively small change in model structure can cause large declines in forecast accuracy. When fit using data from one or two species, the LV3 model, which differs from the HP model only in the form of the functional response, had an average forecast error of over 40%, which was no better than using the mean of the time series for prediction. The LV3 model outperformed the other incorrect models when fit using data from three species, possibly as a result of the structural similarity between the LV3 and HP model. However, even then the LV3 model was only marginally better than predicting the mean of the time series. Our results strongly support prior work (Wood and Thomas 1999, Skalski and Gilliam 2001) showing that parametric models that deviate even slightly from the correct structure may provide very poor forecasts.

Additionally, the inclusion of process error to the LV5 model did not improve forecast accuracy relative to the other incorrect models (Fig. 3). Gelman-Rubin diagnostics suggested convergence; however, estimates of observation error were often significantly higher than that used to generate the data (Appendix: Fig. A1). This was due to the tendency of the best-fit LV5 parameterizations to be those that yielded steady-state dynamics, with excess variance absorbed by the error terms. Ultimately, the forecast accuracy of the LV5 model was similar to that of the LV3 model, suggesting that model misspecification is not alleviated by the addition of process error. Alternative goodness-of-fit criteria do exist (e.g., Lyapunov exponents; Rowlands and Sprott 1992); future work should evaluate their robustness to model misspecification.

Considering the recent push towards ecosystem-based management, an important question is whether the forecast accuracy of models will benefit from including the time series of additional species. We found that the inclusion of additional species yielded a large improvement for the nonparametric model, a smaller improvement for the HP model, and only a slight improvement for the LV3 model. For the MS-Map, additional species acted primarily as a noise-reduction mechanism as the model was better able to determine nearest neighbor points in the state-space reconstruction of the target species. Since the observation errors were assumed to be independent, points which were spuriously close to the predictand (the point to be predicted) due to noise contamination in the one-species map were less likely to be weighted heavily in the three-species map. This explains why, in the absence of noise, there was only a small improvement in MS-Map forecasts when addi-

tional species were added, in contrast with the much larger improvement under the high-noise scenario.

Since the MS-Map was always better than the wrong model, and often better than the correct model in the presence of noise, we suggest that, in real ecosystems, where the true model is never known, a flexible nonparametric framework may be the best method for converting multispecies time series into forecasts.

Several future improvements can be made to the MS-Map. First, a spatially explicit version could provide large gains in forecast accuracy (e.g., Glaser et al. 2011). For each species, both a mean location and abundance could be used to form an embedding. Then, the MS-Map could be used analogously to forecast both location and abundance. Second, the MS-Map could easily be extended to include the magnitude and location of relevant physical variables such as seasonal temperature anomalies or an ocean oscillation index (Bertignac et al. 1998).

Given that real ecosystem dynamics are highly uncertain, and that nonparametric forecasting appears to be robust to such uncertainty, future work should test whether nonparametrics and short-term forecasts can improve resource management. Building on previous work (e.g., Ott et al. 1990, Desharnais et al. 2001), new conservation strategies based on attractor reconstruction could be developed. For example, one could design a harvest rule that minimizes the probability that a species will enter a high-risk area of phase space using short-term predictions. Evaluating these ideas and comparing them to more traditional assessment methods should be a priority, particularly as we move toward implementing ecosystem approaches to management.

#### ACKNOWLEDGMENTS

We gratefully acknowledge the helpful comments of P. de Valpine, E. Pikitch, E. Deyle, H. Ye, and three anonymous reviewers. Funding support was provided by a NMFS/Sea Grant Population Dynamics Fellowship (E/PD-9), grant NA08OAR4320894 from the CAMEO program (a partnership between the NSF and NOAA), and NSF grant DEB 1020372.

#### LITERATURE CITED

- Berliner, L. M. 1991. Likelihood and Bayesian prediction of chaotic systems. *Journal of the American Statistical Association* 86:938–952.
- Bertignac, M., P. Lehodey, and J. Hampton. 1998. A spatial population dynamics simulation model of tropical tunas using a habitat index based on environmental parameters. *Fisheries Oceanography* 7:326–334.
- Cao, L., A. Mees, and K. Judd. 1998. Dynamics from multivariate time series. *Physica D: Nonlinear Phenomena* 121:75–88.
- Christensen, V., and C. J. Walters. 2004. Ecopath with Ecosim: methods, capabilities and limitations. *Ecological Modelling* 172:109–139.
- Clark, J. S., and O. N. Bjørnstad. 2004. Population time series: process variability, observation errors, missing values, lags, and hidden states. *Ecology* 85:3140–3150.
- Cressie, N., C. A. Calder, J. S. Clark, J. M. Ver Hoef, and C. K. Wikle. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling. *Ecological Applications* 19:553–570.
- Desharnais, R. A., R. F. Costantino, J. M. Cushing, S. M. Henson, and B. Dennis. 2001. Chaos and population control of insect outbreaks. *Ecology Letters* 4:229–235.
- Deyle, E. R., and G. Sugihara. 2011. Generalized theorems for nonlinear state space reconstruction. *PLoS ONE* 6:e18295.
- Dixon, P. A., M. J. Milicich, and G. Sugihara. 1999. Episodic fluctuations in larval supply. *Science* 283:1528–1530.
- Fulton, E. A., J. S. Link, I. C. Kaplan, M. Savina Rolland, P. Johnson, C. Ainsworth, P. Horne, R. Gorton, R. J. Gamble, A. D. M. Smith, and D. C. Smith. 2011. Lessons in modelling and management of marine ecosystems: the Atlantis experience. *Fish and Fisheries* 12:171–188.
- Fulton, E. A., A. D. Smith, and A. E. Punt. 2005. Which ecological indicators can robustly detect effects of fishing? *ICES Journal of Marine Science* 62:540–551.
- Geyer, C. 1991. Markov chain Monte Carlo maximum likelihood. Pages 156–163 in E. M. Keramidas and S. M. Kaufman, editors. *Computing science and statistics. Proceedings of the 23rd Symposium on the Interface*. Interface Foundation of North America, Seattle, Washington, USA.
- Glaser, S. M., H. Ye, M. Maunder, A. MacCall, M. Fogarty, and G. Sugihara. 2011. Detecting and forecasting complex nonlinear dynamics in spatially structured catch-per-unit-effort time series for North Pacific albacore (*Thunnus alalunga*). *Canadian Journal of Fisheries and Aquatic Sciences* 68:400–412.
- Grumbine, R. E. 1994. What is ecosystem management? *Conservation Biology* 8:27–38.
- Hardle, W., H. Lutkepohl, and R. Chen. 1997. A review of nonparametric time series analysis. *International Statistical Review* 65:49–72.
- Hastings, A., and T. Powell. 1991. Chaos in a three-species food chain. *Ecology* 72:896–903.
- Holmes, E. E., J. L. Sabo, S. V. Viscido, and W. F. Fagan. 2007. A statistical approach to quasiextinction forecasting. *Ecology Letters* 10:1182–1198.
- Hsieh, C.-H., C. Anderson, and G. Sugihara. 2008. Extending nonlinear analysis to short ecological time series. *American Naturalist* 171:71–80.
- Hsieh, C.-H., S. M. Glaser, A. J. Lucas, and G. Sugihara. 2005. Distinguishing random environmental fluctuations from ecological catastrophes for the North Pacific Ocean. *Nature* 435:336–340.
- Ives, A. R., K. C. Abbott, and N. L. Ziebarth. 2010. Analysis of ecological time series with ARMA(p,q) models. *Ecology* 91:858–871.
- Ives, A. R., K. Gross, and J. L. Klug. 1999. Stability and variability in competitive communities. *Science* 286:542–544.
- Ludwig, D., and C. J. Walters. 1985. Are age-structured models appropriate for catch-effort data? *Canadian Journal of Fisheries and Aquatic Sciences* 42:1066–1072.
- Ludwig, D., C. J. Walters, and J. Cooke. 1988. Comparison of two models and two estimation methods for catch and effort data. *Natural Resource Modeling* 2:457–498.
- MathWorks. 2008. MATLAB version 7.7.0. MathWorks, Natick, Massachusetts, USA.
- Ott, E., C. Grebogi, and J. A. Yorke. 1990. Controlling chaos. *Physical Review Letters* 64:1196.
- Pikitch, E. K., et al. 2004. Ecosystem-based fishery management. *Science* 305:346–347.
- Plummer, M. 2003. JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. Pennsylvania State University, University Park, Pennsylvania, USA. <http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.13.3406>
- Polansky, L., P. de Valpine, J. O. Lloyd Smith, and W. M. Getz. 2009. Likelihood ridges and multimodality in population growth rate models. *Ecology* 90:2313–2320.
- R Development Core Team. 2011. R: a language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. <http://www.R-project.org/>

- Reed, D. H., and G. R. Hobbs. 2004. The relationship between population size and temporal variability in population size. *Animal Conservation* 7:1–8.
- Rowlands, G., and J. Sprott. 1992. Extraction of dynamical equations from chaotic data. *Physica D: Nonlinear Phenomena* 58:251–259.
- Schaefer, M. 1957. Some considerations of population dynamics and economics in relation to the management of the commercial marine fisheries. *Journal of the Fisheries Research Board of Canada* 14:669–681.
- Schreiber, T. 1999. Interdisciplinary application of nonlinear time series methods. *Physics Reports* 308:1–64.
- Skalski, G. T., and J. F. Gilliam. 2001. Functional responses with predator interference: viable alternatives to the Holling Type II model. *Ecology* 82:3083–3092.
- Slocombe, D. S. 1998. Defining goals and criteria for ecosystem-based management. *Environmental Management* 22:483–493.
- Steele, J. H., and J. J. Ruzicka. 2011. Constructing end-to-end models using ECOPATH data. *Journal of Marine Systems* 87:227–238.
- Sugihara, G. 1994. Nonlinear forecasting for the classification of natural time series. *Philosophical Transactions: Physical Sciences and Engineering* 348:477–495.
- Sugihara, G., M. Casdagli, E. Habjan, D. Hess, P. Dixon, and G. Holland. 1999. Residual delay maps unveil global patterns of atmospheric nonlinearity and produce improved local forecasts. *Proceedings of the National Academy of Sciences USA* 96:14210–14215.
- Sugihara, G., B. Grenfell, and R. M. May. 1990. Distinguishing error from chaos in ecological time series. *Philosophical Transactions of the Royal Society of London B* 330:235–251.
- Sugihara, G., and R. M. May. 1990. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. *Nature* 344:734–741.
- Takens, F. 1981. Detecting strange attractors in turbulence. Pages 366–381 in D. Rand and L.-S. Young, editors. *Dynamical systems and turbulence*, Warwick 1980. Springer, Coventry, UK.
- Takeuchi, Y. 1996. Global dynamical properties of Lotka-Volterra systems. World Scientific, Singapore.
- Wood, S. N., and M. B. Thomas. 1999. Supersensitivity to structure in biological models. *Proceedings of the Royal Society of London B* 266:565–570.

## SUPPLEMENTAL MATERIAL

### Appendix

Tables and figures describing parameter values used for simulations, convergence diagnostics, and forecast performance of all methods ([Ecological Archives E094-068-A1](#)).

### Supplement

MATLAB function for generating predictions using the state-space reconstruction (SSR) forecasting method ([Ecological Archives E094-068-S1](#)).